

#4



OIPE

## RAW SEQUENCE LISTING

DATE: 02/10/2002

PATENT APPLICATION: US/09/943,689A

TIME: 13:27:42

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\02102002\I943689A.raw

ENTERED

```

4 <110> APPLICANT: David Michalovich
5     Matthew Alan Sims
6     Narjis Shaikh
8 <120> TITLE OF INVENTION: NOVEL COMPOUNDS
11 <130> FILE REFERENCE: GP-30088-D1
13 <140> CURRENT APPLICATION NUMBER: 09/943,689A
14 <141> CURRENT FILING DATE: 2001-08-31
16 <150> PRIOR APPLICATION NUMBER: US 09/184,001
17 <151> PRIOR FILING DATE: 1998-11-02
19 <150> PRIOR APPLICATION NUMBER: UK 9806221.9
20 <151> PRIOR FILING DATE: 1998-03-23
22 <150> PRIOR APPLICATION NUMBER: UK 9817479.0
23 <151> PRIOR FILING DATE: 1998-08-11
25 <160> NUMBER OF SEQ ID NOS: 4
27 <170> SOFTWARE: FastSEQ for Windows Version 3.0
29 <210> SEQ ID NO: 1
30 <211> LENGTH: 2186
31 <212> TYPE: DNA
32 <213> ORGANISM: HOMO SAPIENS
34 <400> SEQUENCE: 1
35   cttggcacga ggcctcgtgc caggcggcat gagggcccgc ggcccggggg gctgaggcgc      60
36   ccgccgcctg ccgcgggggc cgctcgcgtc ctocatggag gccggagagg aaccgcgtgct      120
37   gctggccgaa ctcaagcccg ggcgcccca ccagtttgat tggaagtcca gctgtgaaac      180
38   ctggagcgtc gccttctccc agatggctc ctggtttgct tgggtctcaag gacactgcat      240
39   cgtcaaactg atcccctggc cgttgaggga gcagttcatc cctaaagggg ttgaagccaa      300
40   aagccgaagt agcaaaaatg agacgaaagg gcggggcagc ccaaaagaga agacgctgga      360
41   ctgtggtcag attgtctggg ggctggcctt cagcccggtg ccttccccac ccagcaggaa      420
42   gctctgggca cgccaccacc cccaagtgcc cgatgtctct tgcctggttc ttgctacggg      480
43   actcaacgat gggcagatca agatctggga ggtgcagaca gggtcctgc ttttgaatct      540
44   ttccggccac caagatgtcg tgagagatct gagcttcaca cccagtggca gtttgatttt      600
45   ggtctcccg ctcacgggata agactcttcg catctgggac ctgaataaac acggtaaaca      660
46   gattcaagtg ttatcgggac acctgcagtg ggtttactgc tgttccatct cccagactg      720
47   cagcatgctg tgcctcgcag ctggagagaa gtcgggtctt ctatggagca tgaggctcta      780
48   cacgttaatt cggaagctag agggccatca aagcagtgtt gtctcttggt acttctcccc      840
49   cgactctgcc ctgcttggtc cggcttctta cgataccaat gtgattatgt gggacccta      900
50   caccggcgaa aggetgaggt cactccacca caccagggtt gaccccgcta tggatgacag      960
51   tgacgtccac attagctcac tgagatctgt gtgcttctct ccagaaggct tgtacctgac      1020
52   cacggtggca gatgacagac tcctcaggat ctggggccctg gaactgaaaa ctcccattgc      1080
53   atttgctcct atgaccaatg ggcttgctg caccattttt ccacatggtg gagtcattgc      1140
54   cacagggaca agagatggcc acgtccagtt ctggacagct cctaggggtc tgcctcact      1200
55   gaagcactta tgccggaaa ccttcgaag ttctctaaca acttaccag tcctagcact      1260
56   gccaatcccc aagaaaatga aagagttcct cacatacagg actttttaag caacaccaca      1320
57   tcttgtgctt ctttgtagca gggtaaactc tcctgtcaaa gggagttgct ggaataatgg      1380

```

## RAW SEQUENCE LISTING

DATE: 02/10/2002

PATENT APPLICATION: US/09/943,689A

TIME: 13:27:43

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\02102002\I943689A.raw

```

58 gccaaacatc tgggtcttgca ttgaaatagc atttctttgg gattgtgaat agaatgtagc 1440
59 aaaaccagat tccagtgtac tagtcatgga tctttctctc cctggcatgt gaaagtcagt 1500
60 cttagaggaa gagattccac ttgcacggca acagagcctt acgttaaatt ttcagtcag 1560
61 ttatgaacag caagtgttga actctttctg cttgttttga ttcaaagtgc agttactgat 1620
62 gttgttttga ttatgcaact aagtaggcct ccagagcctc tctagtggca gagcagctca 1680
63 cactccctcc gctgggaacg atggcttctg cctagtacct atccttgtgt ttctgatgca 1740
64 gtggttagcat tggttcaagt tctctcctgc tgtggtcaga gttgcttcga tgttgccaa 1800
65 gtgcttttct tcttgggctc ccttctgacc tgcaggacag ttttcttgga gccatttgg 1860
66 atgaggtatt aatttagctt aactaaatta caggggactc agaggccgtg ctctgaccg 1920
67 atccagacac tattactggc tttttttttt ttttttttaa caatggtgtg catgtgcagg 1980
68 aaatgacaaa tttgtatgtc agattataca aggatgtatt cttaaaccgc atgactattc 2040
69 agatggctac tgagttatca gtggccattt attagcatca tatttatttg tattttctca 2100
70 acagatgtta aggtacaact gtgttttctt cgattatcta aaaaccatag tacttaaatt 2160
71 gaaaaaaaaa aaaaaaaaaa aaaaaa 2186
73 <210> SEQ ID NO: 2
74 <211> LENGTH: 404
75 <212> TYPE: PRT
76 <213> ORGANISM: HOMO SAPIENS
78 <400> SEQUENCE: 2
79 Met Glu Ala Gly Glu Glu Pro Leu Leu Leu Ala Glu Leu Lys Pro Gly
80 1 5 10 15
81 Arg Pro His Gln Phe Asp Trp Lys Ser Ser Cys Glu Thr Trp Ser Val
82 20 25 30
83 Ala Phe Ser Pro Asp Gly Ser Trp Phe Ala Trp Ser Gln Gly His Cys
84 35 40 45
85 Ile Val Lys Leu Ile Pro Trp Pro Leu Glu Glu Gln Phe Ile Pro Lys
86 50 55 60
87 Gly Phe Glu Ala Lys Ser Arg Ser Ser Lys Asn Glu Thr Lys Gly Arg
88 65 70 75 80
89 Gly Ser Pro Lys Glu Lys Thr Leu Asp Cys Gly Gln Ile Val Trp Gly
90 85 90 95
91 Leu Ala Phe Ser Pro Trp Pro Ser Pro Ser Arg Lys Leu Trp Ala
92 100 105 110
93 Arg His His Pro Gln Val Pro Asp Val Ser Cys Leu Val Leu Ala Thr
94 115 120 125
95 Gly Leu Asn Asp Gly Gln Ile Lys Ile Trp Glu Val Gln Thr Gly Leu
96 130 135 140
97 Leu Leu Leu Asn Leu Ser Gly His Gln Asp Val Val Arg Asp Leu Ser
98 145 150 155 160
99 Phe Thr Pro Ser Gly Ser Leu Ile Leu Val Ser Ala Ser Arg Asp Lys
100 165 170 175
101 Thr Leu Arg Ile Trp Asp Leu Asn Lys His Gly Lys Gln Ile Gln Val
102 180 185 190
103 Leu Ser Gly His Leu Gln Trp Val Tyr Cys Cys Ser Ile Ser Pro Asp
104 195 200 205
105 Cys Ser Met Leu Cys Ser Ala Ala Gly Glu Lys Ser Val Phe Leu Trp
106 210 215 220
107 Ser Met Arg Ser Tyr Thr Leu Ile Arg Lys Leu Glu Gly His Gln Ser
108 225 230 235 240

```

## RAW SEQUENCE LISTING

DATE: 02/10/2002

PATENT APPLICATION: US/09/943,689A

TIME: 13:27:43

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\02102002\I943689A.raw

```

109 Ser Val Val Ser Cys Asp Phe Ser Pro Asp Ser Ala Leu Leu Val Thr
110                245                250                255
111 Ala Ser Tyr Asp Thr Asn Val Ile Met Trp Asp Pro Tyr Thr Gly Glu
112                260                265                270
113 Arg Leu Arg Ser Leu His His Thr Gln Val Asp Pro Ala Met Asp Asp
114                275                280                285
115 Ser Asp Val His Ile Ser Ser Leu Arg Ser Val Cys Phe Ser Pro Glu
116                290                295                300
117 Gly Leu Tyr Leu Ala Thr Val Ala Asp Asp Arg Leu Leu Arg Ile Trp
118                305                310                315                320
119 Ala Leu Glu Leu Lys Thr Pro Ile Ala Phe Ala Pro Met Thr Asn Gly
120                325                330                335
121 Leu Cys Cys Thr Phe Phe Pro His Gly Gly Val Ile Ala Thr Gly Thr
122                340                345                350
123 Arg Asp Gly His Val Gln Phe Trp Thr Ala Pro Arg Val Leu Ser Ser
124                355                360                365
125 Leu Lys His Leu Cys Arg Lys Ala Leu Arg Ser Phe Leu Thr Thr Tyr
126                370                375                380
127 Gln Val Leu Ala Leu Pro Ile Pro Lys Lys Met Lys Glu Phe Leu Thr
128                385                390                395                400
129 Tyr Arg Thr Phe

```

132 &lt;210&gt; SEQ ID NO: 3

133 &lt;211&gt; LENGTH: 2558

134 &lt;212&gt; TYPE: DNA

135 &lt;213&gt; ORGANISM: HOMO SAPIENS

W--&gt; 136 &lt;220&gt; FEATURE:

137 &lt;221&gt; NAME/KEY: UNSURE

138 &lt;222&gt; LOCATION: (146)(161)(178)(233)(254)(296)

140 &lt;400&gt; SEQUENCE: 3

```

141 gaaaggaacc gctgctgctg gccgaactca agcccgggcg cccccaccag tttgattgga      60
142 agtcagctg tgaaacctgg agcgtcgctt tctccccaga tggctcctgg tttgcttggt      120
W- 143 ctcaaggaca ctgcatcgtc aaactnatcc cctggccggt ngaggagcag ttcattccta      180
W- 144 aaggggttga agccaaaagc cgaagtagca aaaatgagac gaaagggcgg ggnagtccaa      240
W- 145 aagagaagac gctngactgt ggtcagattg tctggggggt ggccttcagc ctgtgncitt      300
146 cccacccag caggaagctc tgggcacgcc accaccccca agtgccgat gtctcttgcc      360
147 tggttcttgc tacgggactc cacgatgggc agatcaagat ctgggaggtg cagacagggc      420
148 tctgctttt gaatctttcc ggccaccaag atgtcgtgag agatctgagc ttcacacca      480
149 gtggcagttt gatttttggtc tccgcgtcac gggataagac tcttcgcata tgggacctga      540
150 ataaacacgg taaacagatt caagtgttat cgggccacct gcagtgggtt tactgctgtt      600
151 ccatctcccc agactgcagc atgctgtgct ctgcagctgg agagaagtcg gtctttctat      660
152 ggagcatgag gtcctacacg ttaattcgga agctagaggg ccatcaaagc agtggtgtct      720
153 cttgtgactt ctcccccgac tctgcctgc ttgtcacggc ttcttacgat accaatgtga      780
154 ttatgtggga cccctacacc ggcgaaaggc tgaggtcact ccaccacacc caggttgacc      840
155 ccgccatgga tgacagtgac gtccacatta gctcactgag atctgtgtgc ttctctccag      900
156 aaggcttgta ccttgccacg gtggcagatg acagactcct caggatctgg gccctggaac      960
157 tgaaaactcc cattgcattt gtcctatga ccaatgggct ttgctgcaca ttttttccac      1020
158 atggtggagt cattgcaca gggacaagag atggccacgt ccagttctgg acagctocta      1080
159 gggtcctgtc ctcactgaag cacttatgcc ggaaagccct tcgaagtttc ctaacaactt      1140
160 accaagtcct agcactgcca atccccaaaga aaatgaaaga gttcctcaca tacaggactt      1200

```

## RAW SEQUENCE LISTING

DATE: 02/10/2002

PATENT APPLICATION: US/09/943,689A

TIME: 13:27:43

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\02102002\I943689A.raw

```

161 ttttaagcaac accacatctt gtgcttcttt gtagcagggt aaatcgtcct gtcaaaggga 1260
162 gttgctggaa taatgggcca aacatctggt cttgcattga aatagcattt ctttgggatt 1320
163 gtgaatagaa ttagcaaaa ccagattcca gtgtactagt catggatctt tctctccctg 1380
164 gcatgtgaaa gtcagcttta gaggaagaga ttccacttgc acggcaacag agccttacgt 1440
165 taaattttca gtccagttat gaacagcaag tgttgaactc tttctgcttg ttttgattca 1500
166 aagtgcagtt actgatgttg ttttgattat gcaactaagt aggcctccag agcctctcta 1560
167 gtggcagagc agctcacact ccctccgctg ggaacgatgg cttctgccta gtacctatcc 1620
168 ttgtgtttct gatgcagtgg tagcattggt tcaagttctc tctgctgtg gtcagagttg 1680
169 cttgatgtt ggccaagtgc ttttcttctt gggctccctt ctgacctgca ggacagtttt 1740
170 cctggagcca tttggtatga ggtattaatt tagcttaact aaattacagg ggactcagag 1800
171 gccgtgctcc tgaccgatcc agacactatt actggctttt tttttttttt ttttaacaatg 1860
172 gtgtgcatgt gcaggaaatg acaaatattgt atgtcagatt atacaaggat gtattcttaa 1920
173 accgcatgac tattcagatg gctactgagt tatcagtggc catttattag catcatattt 1980
174 atttgtattt tctcaacaga tgtaaaggta caactgtgtt tttctcgatt atctaaaaac 2040
175 catagtactt aaattgaaca gttgcaaaga tgtcttaatt gtgtaaagaa ttggtgtagt 2100
176 catgacttta gctgatactc ttatgtacga gatctgtctc tgctgtttaa cttcattgga 2160
177 ttaatcagct gggttcaact ctactgcgaa acaaaaatag ctcccttaaaa gtactgttct 2220
178 ccttcagtggt catgtagtta tctaatacaag acacctcatt caaacaaaaac ctgccttagg 2280
179 aaaatttaat atattttaaa ttattttaaa agaaatacaa catcttattc tttagctttc 2340
180 ttaatcgggtg ctttatggag gccagtgtaa cgttacatga ctcgttgaga aagttgagga 2400
181 atttctctta ccacctttgt tgcttgaaga aaaacatgtc ttttcaaaat gagaggcttt 2460
182 cattgaagaa aagaaaaaaa caacagttaa aagctattgg ctctctgttt catttttttc 2520
183 cattaagaaa aaaaaaagtc cccttttaaa acaagcag 2558

```

185 &lt;210&gt; SEQ ID NO: 4

186 &lt;211&gt; LENGTH: 400

187 &lt;212&gt; TYPE: PRT

188 &lt;213&gt; ORGANISM: HOMO SAPIENS

W--&gt; 189 &lt;220&gt; FEATURE:

190 &lt;221&gt; NAME/KEY: UNSURE

191 &lt;222&gt; LOCATION: (53)(59)(98)

193 &lt;400&gt; SEQUENCE: 4

194 Lys Glu Pro Leu Leu Leu Ala Glu Leu Lys Pro Gly Arg Pro His Gln

195 1 5 10 15

196 Phe Asp Trp Lys Ser Ser Cys Glu Thr Trp Ser Val Ala Phe Ser Pro

197 20 25 30

198 Asp Gly Ser Trp Phe Ala Trp Ser Gln Gly His Cys Ile Val Lys Leu

199 35 40 45

W--&gt; 200 Ile Pro Trp Pro Xaa Glu Glu Gln Phe Ile Xaa Lys Gly Phe Glu Ala

201 50 55 60

202 Lys Ser Arg Ser Ser Lys Asn Glu Thr Lys Gly Arg Gly Ser Pro Lys

203 65 70 75 80

204 Glu Lys Thr Leu Asp Cys Gly Gln Ile Val Trp Gly Leu Ala Phe Ser

205 85 90 95

W--&gt; 206 Leu Xaa Leu Ser Pro Pro Ser Arg Lys Leu Trp Ala Arg His His Pro

207 100 105 110

208 Gln Val Pro Asp Val Ser Cys Leu Val Leu Ala Thr Gly Leu His Asp

209 115 120 125

210 Gly Gln Ile Lys Ile Trp Glu Val Gln Thr Gly Leu Leu Leu Leu Asn

211 130 135 140

## RAW SEQUENCE LISTING

DATE: 02/10/2002

PATENT APPLICATION: US/09/943,689A

TIME: 13:27:43

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\02102002\I943689A.raw

```

212 Leu Ser Gly His Gln Asp Val Val Arg Asp Leu Ser Phe Thr Pro Ser
213 145 150 155 160
214 Gly Ser Leu Ile Leu Val Ser Ala Ser Arg Asp Lys Thr Leu Arg Ile
215 165 170 175
216 Trp Asp Leu Asn Lys His Gly Lys Gln Ile Gln Val Leu Ser Gly His
217 180 185 190
218 Leu Gln Trp Val Tyr Cys Cys Ser Ile Ser Pro Asp Cys Ser Met Leu
219 195 200 205
220 Cys Ser Ala Ala Gly Glu Lys Ser Val Phe Leu Trp Ser Met Arg Ser
221 210 215 220
222 Tyr Thr Leu Ile Arg Lys Leu Glu Gly His Gln Ser Ser Val Val Ser
223 225 230 235 240
224 Cys Asp Phe Ser Pro Asp Ser Ala Leu Leu Val Thr Ala Ser Tyr Asp
225 245 250 255
226 Thr Asn Val Ile Met Trp Asp Pro Tyr Thr Gly Glu Arg Leu Arg Ser
227 260 265 270
228 Leu His His Thr Gln Val Asp Pro Ala Met Asp Asp Ser Asp Val His
229 275 280 285
230 Ile Ser Ser Leu Arg Ser Val Cys Phe Ser Pro Glu Gly Leu Tyr Leu
231 290 295 300
232 Ala Thr Val Ala Asp Asp Arg Leu Leu Arg Ile Trp Ala Leu Glu Leu
233 305 310 315 320
234 Lys Thr Pro Ile Ala Phe Ala Pro Met Thr Asn Gly Leu Cys Cys Thr
235 325 330 335
236 Phe Phe Pro His Gly Gly Val Ile Ala Thr Gly Thr Arg Asp Gly His
237 340 345 350
238 Val Gln Phe Trp Thr Ala Pro Arg Val Leu Ser Ser Leu Lys His Leu
239 355 360 365
240 Cys Arg Lys Ala Leu Arg Ser Phe Leu Thr Thr Tyr Gln Val Leu Ala
241 370 375 380
242 Leu Pro Ile Pro Lys Lys Met Lys Glu Phe Leu Thr Tyr Arg Thr Phe
243 385 390 395 400

```

## VERIFICATION SUMMARY

DATE: 02/10/2002

PATENT APPLICATION: US/09/943,689A

TIME: 13:27:44

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\02102002\I943689A.raw

L:136 M:283 W: Missing Blank Line separator, <220> field identifier  
L:143 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:3  
L:143 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:144 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:3  
L:144 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:145 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:3  
L:145 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:189 M:283 W: Missing Blank Line separator, <220> field identifier  
L:200 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:4  
L:200 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4  
L:206 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:4  
L:206 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4